

# **A MATHEMATICAL MODEL FOR LUNG CANCER: THE EFFECTS OF SECOND-HAND SMOKE AND EDUCATION**

**BU-1525 -M**

**August 1999**

**MTBI**

**Carlos A. Acevedo-Estephania**

**Christina Gonzalez**

**Karen R. Rios-Soto**

**Eric D. Summerville**

**Advisors: Baojun Song, Carlos Castillo-Chavez**

**Keywords:** none provided.

## **Abstract:**

In the United States, lung cancer is the leading cause of cancer deaths. As of today, cigarette smoking causes 85% of lung cancer deaths. In this study, a non-linear system of differential equations is used to model the dynamics of a population, which includes smokers. The parameters of the model are obtained from data published by cancer institutes, health and government organizations. The average number of individuals who become smokers and the reduction of this average by an education program are determined. The long-term impact of educating a susceptible class before they enter the population model and the effect it has on the epidemic (getting lung cancer) is also studied. Simulations using realistic parameters are carried out to illustrate our theoretical results.

BU-1525-M

# A Mathematical Model for Lung Cancer: The Effects of Second-Hand Smoke and Education

Carlos A. Acevedo-Estefanía\*

UNIVERSITY OF PUERTO RICO-CAYEY

Christina González†

TEXAS A&M UNIVERSITY

Karen R. Ríos-Soto‡

UNIVERSITY OF PUERTO RICO-MAYAGÜEZ

Eric D. Summerville§

ST. MARY'S UNIVERSITY

## Advisors

Baojun Song

Carlos Castillo-Chavez

*Department of Biometrics - Cornell University*

August 11, 2000

---

\*redewasu@hotmail.com, †c0g7680@labs.tamu.edu, ‡peronce@yahoo.com, §cire127@hotmail.com

### **Abstract**

In the United States, lung cancer is the leading cause of cancer deaths. As of today, cigarette smoking causes 85 percent of lung cancer deaths. In this study, a non-linear system of differential equations is used to model the dynamics of a population which includes smokers. The parameters of the model are obtained from data published by cancer institutes, health and government organizations. The average number of individuals who become smokers and the reduction of this average by an education program are determined. The long-term impact of educating a susceptible class before they enter the population model and the effect it has on the epidemic is also studied. Simulations using realistic parameters are carried out to illustrate our theoretical results.

# 1 Introduction

Lung cancer, also referred to as bronchogenic carcinomas, is a major contributor of cancer deaths in the United States, accounting for 28 percent of such deaths [8]. The development of lung cancer occurs on the lining glands, which contains damage cells that are located in our lungs and broncheal airways known as the tracheobronchial system [3]. This part of a human being is important because this system is susceptible to being contaminated by inhaled air, which is a major factor in the development of lung cancer. Scientists believe that the major cause of lung cancer is due to cancer-causing agents known as carcinogens, such as asbestos and radon. However, research and statistics show that the major agent of lung cancer is tobacco smoke, which contains over 60 carcinogens.

Today, cigarette smoke is responsible for a great proportion of deaths within tobacco smoke. Each year in the United States, approximately 400,000 people die from cigarette smoke, which accounts for one in every five deaths in the United States [14]. The likelihood that a smoker will develop lung cancer from cigarette smoke depends on many aspects; such as the age at which smoking began, how long the person has smoked, the number of cigarettes smoked per day, and how deeply the smoker inhales [10]. In 1988, the Surgeon General established the addictive potential of cigarette smoking by stressing that nicotine and other agents in cigarettes were just as addictive as cocaine [8]. The ability of a smoker to quit is very difficult because of the addiction to nicotine. In fact, 90 percent of smokers would like to quit but can not [12]. Based on data of current smokers, only 34 percent of smokers attempt to quit, but only 2.5 percent succeed every year [8]. The use of cigarettes and the toxic air it creates has been labeled as the single most preventable cause of premature death in the United States.

The relationship between cigarette smoke with respect to lung cancer has been established in 85-90 percent of all lung cancer cases (146,000 case/year). Furthermore, an estimated 3,000 non-smokers per year die from lung cancer due to second-hand smoke (also known as environmental tobacco smoke, ETS) [14]. The number of deaths of non-smokers may be lower than active smokers, but according to the U.S. Environmental Protection Agency, it is quite large when compared to those associated with other indoor and outdoor environmental pollutants. This data has had a great impact on public policies that protect people from second-hand smoke [9].

Based on the relationship between lung cancer and cigarette smoke, we want to

show the reduction of contact between non-smokers and smokers, and how to decrease the rate in which non-smokers and smokers progress towards lung cancer. The arrangement of seven different classes will assist us to define the total population we want to analyze. However, the best way to detail the transition of each class is to use a mixture of parameters, probabilities, and rates. Based on the behavior of each class, seven non-linear differential equations are created. One of the main purposes of the nonlinear equations is to obtain the equilibrium points. The Jacobean Matrix is used to find the basic reproductive number,  $R_0$ , which represents the rate that people get infected. The role of  $R_0$  is to determine if smoking will die out or increase. Through simulations, the model is analyzed to obtain different situations that produce interesting results among the specific classes. Using real life data, the model is believed to show how the increase of the educated class can lower the probability of being diagnosed with lung cancer.

Our nonlinear differential equation model that focuses on the impact of peer pressure on non-smokers and the progression to lung cancer via first and second-hand smoke. The dynamics of addiction are shown to be governed by a single non-dimensional parameter,  $R_0$ .  $R_0$  denotes the number of secondary addictions generated by the first (small) class of smokers in a population of (mostly) non-smokers. Obviously,  $R_0 > 1$  shows how high the prevalence of addicts to nicotine is. Our analysis then focuses on the role of education at various levels of the progression chain (to lung cancer) in the long-term reduction of lung cancer. Our results show that the most important factor in preventing individuals from becoming smokers is education; while the second most important measure is to convince heavy smokers to quit. Our results partially agree with those recently published by *Ithaca Journal*. However, we disagree on the recommendation of focusing education on smokers. The prevention of smoking is most effective in the long run, if it is focused on non-smokers.

Our paper is organized as follows: section 2, explains the population model, while section 3, explains the analysis of the smoke-free equilibrium, the basic reproductive number, and endemic equilibrium. In section 4, we have an estimation of parameters and numerical solutions; section 5, the conclusion; and section 6, the future work.

## 2 A Population Model for the Risk of Getting Lung Cancer

We divide the total population into two sub-populations, which consists of individuals who have never smoked that respond to prevention education and those who did not. The educated population is denoted by individuals who never become smokers,  $E(t)$ . The less-educated population, is made up of six classes. The non-smoker class,  $N(t)$ , includes the individuals who do not smoke but are susceptible to smoking; the light-smoker class,  $I_1(t)$ , includes those who smoke 15 or less cigarettes per day; and the heavy smokers  $I_2(t)$ . There exists three additional classes in the less-educated population:  $Q(t)$ , the quitter class, consists of individuals who used to smoke but stop permanently;  $S(t)$ , who used to smoke and are likely to smoke again; and the lung cancer class,  $L(t)$ , individuals who have developed lung cancer. We treat the people that smoke as an infected group, in order to show the transmission at which the infection of smoking occurs.

An individual can enter the population in two different ways. One way is proceeding into the educated class,  $E$ , with a probability of  $q$ , or into the non-smoker class,  $N$ , with a probability of  $(1-q)$ . Individuals in all classes may develop lung cancer because of the impact of second-hand smoke.

Individuals in the non-smoker class can become a light smoker ( $I_1$ ) due to lack of education and peer pressure of smokers. Once they become a light smoker, they can not move back to  $N$  or  $E$ . Therefore, a light smoker may become a heavy smoker ( $I_2$ ), or they may stop smoking temporarily ( $S$ ) or permanently ( $Q$ ). We assume that in order for them to become a heavy smoker, they must start off as a light smoker.

Once an individual becomes a heavy smoker, s/he may quit temporarily ( $S$ ), permanently ( $Q$ ), or develop lung cancer. In the  $S$  class, the individuals can either go back to smoking, in which we assume they start off as light smokers; or they can develop lung cancer ( $L$ ). The  $Q$  class represents the number of individuals who stop smoking permanently. However, they have a higher probability of developing lung cancer than a non-smoker.

We let the natural death rate (per capita),  $\mu$ , be the same for all the classes except for the  $L$  class, which is considered to have higher death rate.

Our mathematical model is given by the following non-linear system of ordinary differential equations.

$$\frac{dN}{dt} = (1 - q)\Lambda - \frac{\beta N(I_1 + I_2)}{T} - \mu N, \quad (1)$$

$$\frac{dI_1}{dt} = \frac{((1 - P_n)\beta N + (1 - P_s)\beta S)(I_1 + I_2)}{T} - (\sigma_1 + \gamma_1 + \delta_1 + \mu)I_1, \quad (2)$$

$$\frac{dI_2}{dt} = \gamma_1 I_1 - (\gamma_2 + \delta_2 + \mu)I_2, \quad (3)$$

$$\frac{dQ}{dt} = p_2 \gamma_2 I_2 + p_1 \sigma_1 I_1 - (\delta_q + \mu)Q, \quad (4)$$

$$\frac{dS}{dt} = (1 - p_1)\sigma_1 I_1 + (1 - p_2)\gamma_2 I_2 - \frac{\beta S(I_1 + I_2)}{T} - \mu S, \quad (5)$$

$$\begin{aligned} \frac{dL}{dt} = & \frac{(P_n \beta N + P_s \beta S + \beta_e E)(I_1 + I_2)}{T} + \delta_1 I_1 + \delta_2 I_2 + \delta_q Q \\ & - (\mu + d)L, \end{aligned} \quad (6)$$

$$\frac{dE}{dt} = q\Lambda - \frac{\beta_e E(I_1 + I_2)}{T} - \mu E, \quad (7)$$

$$\text{where } T = E + N + I_1 + I_2 + Q + S + L,$$

the parameters and their expected values are listed in Table 1 and Table 2, respectively.

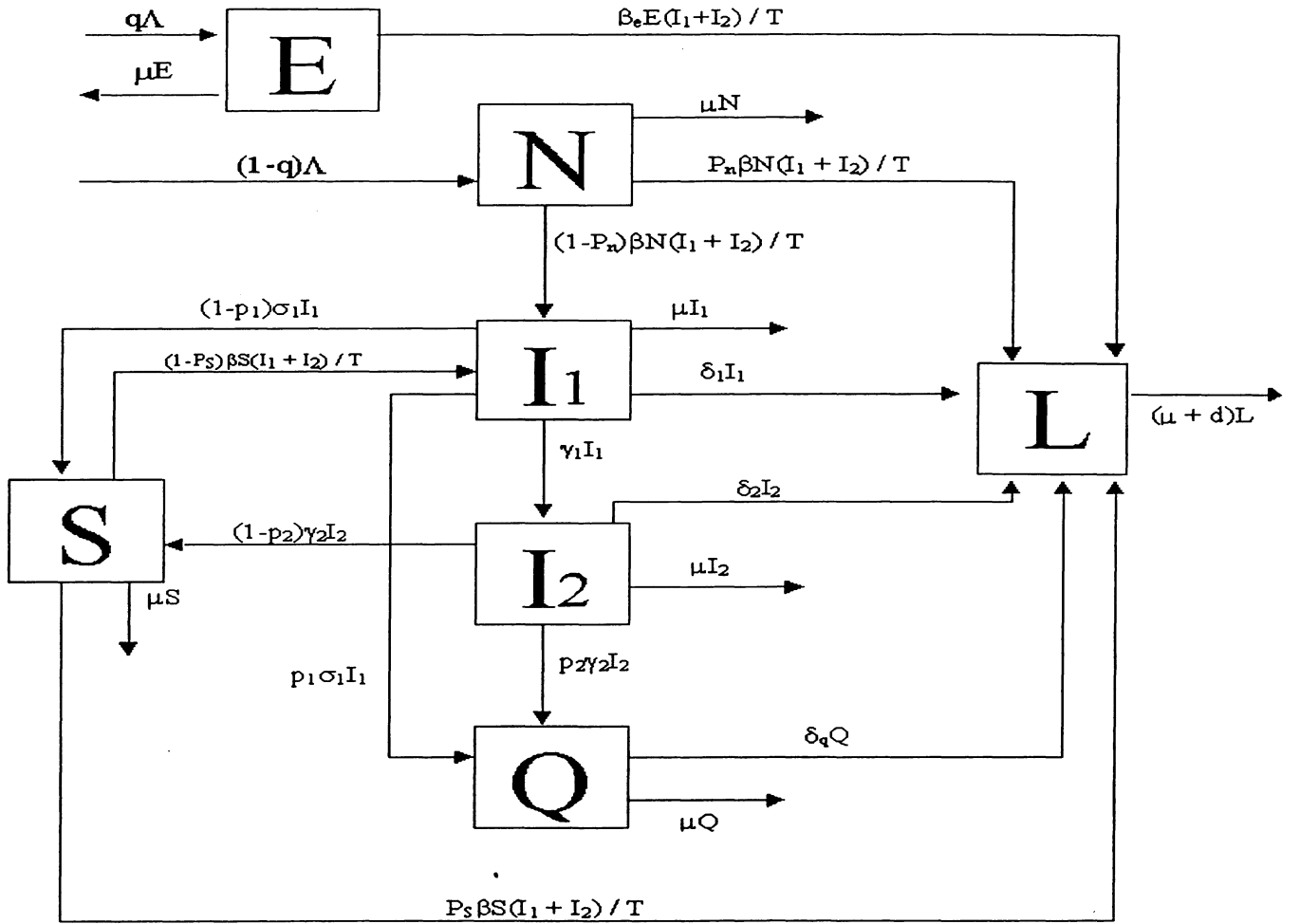


Figure 1: Diagram of the effects of smoking on a population.



Table 1: Table of Parameters

Parameter	Definition
$\Lambda$	Recruitment rate.
$\mu$	Mortality rate (per-capita).
$\beta$	Transmission rate.
$\beta_e$	Rate in which the educated class develops lung cancer due to second-hand smoke.
$\delta_1$	$I_1$ - Rate for developing lung cancer.
$\delta_2$	$I_2$ - Rate for developing lung cancer.
$\delta_q$	$Q$ -Rate for developing lung cancer.
$\gamma_1$	Rate in which light smokers become heavy smokers.
$\gamma_2$	Rate in which a heavy smoker quits smoking.
$\sigma_1$	Rate in which a light smoker stops smoking.
$d$	Mortality rate in which a person dies of lung cancer.
$q$	Probability that an incoming individual enters into the educated class.
$(1-q)$	Probability that a non-educated individual enters the non-smoker class.
$P_n$	Probability that a non-smoker develops lung cancer.
$(1-P_n)$	Probability that a non-smoker becomes a light smoker.
$P_s$	Probability of getting lung cancer via secondary smoke, if you go in $S$ .
$(1-P_s)$	Probability in which a person who stopped smoking temporarily becomes a light smoker.
$p_1$	Probability that a light smoker quits smoking permanently.
$(1-p_1)$	Probability that a light smoker quits smoking temporarily.
$p_2$	Probability that a heavy smoker quits smoking permanently.
$(1-p_2)$	Probability that a heavy smoker quits smoking temporarily.

### 3 Analysis

#### 3.1 Smoke-Free Equilibrium and the Basic Reproductive Number

In this section we analyze the non-linear differential equation model. We solve the system of non-linear differential equations to find the equilibrium points, with the assistance of the mathematical program Maple 6. First, we solve for the smoke-free equilibrium, which is:

$$\left( \frac{\Lambda(1-q)}{\mu}, 0, 0, 0, 0, 0, \frac{q\Lambda}{\mu} \right).$$

Throughout of this paper we consistantly use  $\Sigma_1$  an  $\Sigma_2$  which are:

$$\Sigma_1 = \gamma_1 + \sigma_1 + \delta_1 + \mu \text{ and } \Sigma_2 = \gamma_2 + \delta_2 + \mu$$

The Jacobian Matrix at the smoke-free equilibrium is:

$$\begin{pmatrix} -\mu & -\beta(1-q) & -\beta(1-q) & 0 & 0 & 0 & 0 \\ 0 & (1-P_n)\beta(1-q) - \Sigma_1 & (1-P_n)\beta(1-q) & 0 & 0 & 0 & 0 \\ 0 & \gamma_1 & -\Sigma_2 & 0 & 0 & 0 & 0 \\ 0 & p_1\sigma_1 & p_2\gamma_2 & -(\delta_q + \mu) & 0 & 0 & 0 \\ 0 & (1-p_1)\sigma_1 & (1-p_2)\gamma_2 & 0 & -\mu & 0 & 0 \\ 0 & P_n\beta(1-q) + \beta_e q + \delta_1 & P_n\beta(1-q) + \beta_e q + \delta_2 & \delta_q & 0 & -(\mu + d) & 0 \\ 0 & -\beta_e q & -\beta_e q & 0 & 0 & 0 & -\mu \end{pmatrix}$$

This matrix has 5 negative eigenvalues, which are:

$$-\mu, -(\delta_q + \mu), -\mu, -(\mu + d) \text{ } -\mu.$$

The rest of the eigenvalues are from the sub-matrix:

$$\begin{pmatrix} (1-P_n)\beta(1-q) - \Sigma_1 & (1-P_n)\beta(1-q) \\ \gamma_1 & -\Sigma_2 \end{pmatrix}$$

Local asymptotical stability is guaranteed provided that the determinant is greater than zero, that is, if

$$((1 - P_n)\beta(1 - q) - \Sigma_1)(-\Sigma_2) - ((1 - P_n)\beta(1 - q))\gamma_1 > 0, \quad (8)$$

which is equivalent to

$$1 > \frac{(1 - q)(1 - P_n)\beta(\gamma_1 + \Sigma_2)}{(\Sigma_2)(\Sigma_1)}. \quad (9)$$

Hence, we define:

$$R_0 = \frac{(1 - q)(1 - P_n)\beta(\gamma_1 + \Sigma_2)}{(\Sigma_2)(\Sigma_1)}, \quad (10)$$

and conclude that if  $R_0 < 1$ , then the smoke-free equilibrium point is locally asymptotically stable.  $R_0$  implies a smoke-free population. Note that  $R_0$  can be rewritten as:

$$R_0 = \left( \frac{(1 - q)(1 - P_n)\beta}{\Sigma_1} \right) + \left( \frac{\gamma_1}{\Sigma_1} \right) \left( \frac{(1 - q)(1 - P_n)\beta}{\Sigma_2} \right) \quad (11)$$

Hence, the basic reproductive number,  $R_0$ , gives the number of the secondary smokers produced by a typical smoker during his life as a smoker.

Observe that  $\frac{1}{\Sigma_1}$  is the average amount of time a person stays in the light smoker class ( $I_1$ );  $\frac{1}{\Sigma_2}$  is the average amount of time that a person stays in the heavy smoker class ( $I_2$ );  $(1 - P_n)\beta$  is the rate in which a nonsmoker become a light smoker per unit of time; and,  $(1 - q)$  represent the probability of a non-educated person entering the non-smoker class. Hence,  $\frac{(1 - q)(1 - P_n)\beta}{\Sigma_1}$  represents the new smokers from light smokers;  $\frac{\gamma_1}{\Sigma_1}$  is the proportion of a light smokers who become from heavy smokers; while  $\left( \frac{\gamma_1}{\Sigma_1} \right) \left( \frac{(1 - q)(1 - P_n)\beta}{\Sigma_2} \right)$  represents new smokers from heavy smokers.  $R_0 < 1$  implies a non-smoker society.

Looking at the  $R_0$ , we can analyze the sensitivity of the system by observing the parameters that can drastically change the value of  $R_0$ . The value of  $q$ , which is the probability of getting into the educated class, would have an important effect, particularly if it is closer to one. If we make it approximately equal to one or close enough,

we get the  $R_0$  to be less than one, that is, our population becomes smoke-free. If  $q$  is close to zero, then most of the population will go to the  $N$  class.

Other parameter that greatly affect  $R_0$  is  $\beta$ , since this is the transmission rate between classes. It is obvious by looking of  $R_0$  that if we increase the amount of  $\beta$ ,  $R_0$  will increase, reducing the contact with smokers.

The other parameter that seems important is  $P_n$ ; however  $P_n$ , is very small. In the case of  $R_0 < 1$ , increasing  $P_n$  leads to the increase of people developing lung cancer. As  $t$  gets larger, the number of lung cancer cases goes to zero.

### 3.2 Endemic Equilibria

The previous subsection shows that if  $R_0 < 1$ , then the smoker-free equilibrium is locally asymptotically stable, meaning eventually that there will be non-smokers. To look at the case  $R_0 > 1$ , we solve the following algebraic equations in order to find out whether or not a positive equilibrium is possible.

$$0 = (1 - q)\Lambda - \frac{\beta N(I_1 + I_2)}{T} - \mu N, \quad (12)$$

$$0 = \frac{((1 - P_n)\beta N + (1 - P_s)\beta S)(I_1 + I_2)}{T} - (\sigma_1 + \gamma_1 + \delta_1 + \mu)I_1, \quad (13)$$

$$0 = \gamma_1 I_1 - (\gamma_2 + \delta_2 + \mu)I_2, \quad (14)$$

$$0 = p_2 \gamma_2 I_2 + p_1 \sigma_1 I_1 - (\delta_q + \mu)Q, \quad (15)$$

$$0 = (1 - p_1)\sigma_1 I_1 + (1 - p_2)\gamma_2 I_2 - \frac{\beta S(I_1 + I_2)}{T} - \mu S, \quad (16)$$

$$0 = \frac{(P_n \beta N + P_s \beta S + \beta_e E)(I_1 + I_2)}{T} + \delta_1 I_1 + \delta_2 I_2 + \delta_q Q - (\mu + d)L, \quad (17)$$

$$0 = q\Lambda - \frac{\beta_e E(I_1 + I_2)}{T} - \mu E, \quad (18)$$

where  $T = E + N + I_1 + I_2 + Q + S + L$ .

If we let:

$$A = \frac{\gamma_1}{\Sigma_2}, B = \frac{p_1\gamma_1+p_2\gamma_2A}{\delta_q+\mu}, \Omega = (1-p_n)\sigma_1 + (1-p_2)(\gamma_2A)$$

$$\Phi(I_1) = (1-P_n)(1-q)\Lambda + (1-P_s)I_1(1-p_1)\sigma_1 + (1-p_2)(\gamma_2A).$$

Then using (14) and (15), we can represent  $I_2$  and  $Q$  in terms of  $I_1$ , respectively,

$$I_2 = AI_1 \text{ and} \quad (19)$$

$$Q = BI_1. \quad (20)$$

Multiplying equation (12) by  $S$  and equation (16) by  $N$ , we find a linear relationship between  $S$  and  $N$ , namely

$$S = \frac{\Omega I_1 N}{(1-q)\Lambda}. \quad (21)$$

Using equations (13) and (21), we can solve for  $\frac{N}{T}$ ,

$$\frac{N}{T} = \left( \frac{\Sigma_1}{\beta(1+A)} \right) \left( \frac{(1-q)\Lambda}{\Phi(I_1)} \right). \quad (22)$$

Using equations (12) and (22), we can solve for  $N$ ,

$$N = \frac{(1-q)\Lambda}{\mu} - \left( \frac{\Sigma_1}{\mu} \right) \left( \frac{I_1(1-q)\Lambda}{\Phi(I_1)} \right). \quad (23)$$

By using equation 22, we solve for  $T$  explicitly in terms of  $I_1$ ,

$$T = \left( \frac{\beta(1+A)}{\Sigma_1} \right) \left( \frac{\Phi(I_1) - \Sigma_1 I_1}{\mu} \right). \quad (24)$$

To solve for  $E$  and  $\frac{E}{T}$ , we use equation (18),

$$E = \frac{q\Lambda T}{\beta_e I_1(1+A) + \mu T}, \text{ and} \quad (25)$$

$$\frac{E}{T} = \frac{q\Lambda}{\beta_e I_1(1+A) + \mu T} \quad (26)$$

And using equations (21) and (22), we solve for  $\frac{S}{T}$ ,

$$\frac{S}{T} = \frac{\Sigma_1 I_1 \Omega}{\beta(1+A)\Phi(I_1)} \quad (27)$$

Adding equations (12) through (18), allows us to solve for L in terms of T,

$$L = \frac{\Lambda - \mu T}{d}. \quad (28)$$

Substituting equations (19) through (28) into (17), we have an equilibrium equation for  $I_1$ :

$$F(I_1) = \frac{P_n \Sigma_1 (1-q) \Lambda I_1}{\Phi(I_1)} + \frac{P_s \Sigma_1 \Omega (I_1)^2}{\Phi(I_1)} + \frac{\beta_e q \Lambda I_1}{\frac{\beta}{\Sigma_1} (1-P_n)(1-q) \Lambda + [\beta_e - \beta + \frac{\beta}{\Sigma_1} (1-P_s) \Omega] I_1}$$

$$\left( \delta_1 + \delta_2 A + \delta_q B + \frac{\mu+d}{d} \beta (1+A) - \left( \frac{\mu+d}{d} \right) \left( \frac{\beta(1+A)}{\Sigma_1} \right) (1-P_s) \Omega \right) I_1 + \frac{\mu+d}{d} \Lambda (R_0 - 1) = 0$$

Once we solve for  $F(0)$  and  $F(\infty)$ , we obtain:

$$F(0) = \frac{(\mu+d)}{d} \Lambda (R_0 - 1) \text{ and } F(\infty) = -\infty$$

Using the Intermediate Value Theorem (IVT), we obtain that if  $R_0 > 1$ , then if:

$$\delta_1 + \delta_2 A + \delta_q B + \frac{(\mu+d)}{d} \beta (1+A) + \frac{P_s \Sigma_1}{(1-P_s)} < \frac{\mu+d}{d} \frac{\beta(1+A)}{\Sigma_1} (1-P_s) \Omega \text{ and}$$

$$\beta_e - \beta + \frac{\beta}{\Sigma_1} (1-P_s) \Omega > 0$$

or

$$\delta_1 + \delta_2 A + \delta_q B + \frac{(\mu+d)}{d} \beta (1+A) + \frac{P_s \Sigma_1}{(1-P_s)} > \frac{\mu+d}{d} \frac{\beta(1+A)}{\Sigma_1} (1-P_s) \Omega \text{ and}$$

$$\beta_e - \beta + \frac{\beta}{\Sigma_1} (1-P_s) \Omega < 0.$$

This shows that there exists at least an endemic equilibrium solution.

We have shown the existence of an endemic. Which means that the smoking populations are present. It is hard to determine its stability since we do not have the explicit formula for the endemic. Our numerical simulations support our conjecture that this endemic is locally asymptotically stable.

## 4 Estimation of Parameters and Numerical Solutions

### 4.1 Estimated Parameters

We first estimated the parameters by available data, then used Matlab to numerically solve the system of ordinary differential equations (1)–(7).

#### Estimated Values By Data

- $\mu$  - mortality rate, is estimated by the average life span  $\frac{1}{\mu}$ .
- $p_1$  &  $p_2$  - probability given by data that 2.5% of smokers quit permanently [8].
- $\gamma_1$  - probability given by the data that 60% of smokers are in the  $I_2$  class [17].
- $\sigma_1$  - given by the data that individuals in the  $I_1$  class quit at a higher rate [18].
- $\gamma_2$  - given by data that individuals in the  $I_2$  class quit at a lower rate [18].
- $d$  - mortality rate, given by data that people who develop lung cancer have a mortality rate of 7 years less than  $\mu$  [11].

#### Assumed Values

- $\delta_1$  - assuming that 15 out of 1000  $I_1$  individuals develop lung cancer.
- $\delta_2$  - assuming that 30 out of 1000  $I_2$  individuals develop lung cancer.
- $\delta_q$  - assuming that 5 out of 1000  $Q$  individuals develop lung cancer.
- $P_s$  - assuming that the probability of developing lung cancer due to previous smoking or secondary smoking is low.
- $P_n$  - assuming that the probability of developing lung cancer due to secondary smoke is very low.
- $\beta_e$  - assuming that the transmission rate between the  $E$  population and  $I_1$  and  $I_2$  is very low.
- $\Lambda$  - assuming that there is a constant population that enters our model; it has to be greater than 1.

Table 2: Estimation of Parameters

Parameters	Values
$\Lambda$	14+
$\mu$	0.014 *
$\beta$	"2"
$P_n$	0.00001+
$(1-P_n)$	0.99999+
$\delta_1$	0.015*
$\gamma_1$	0.60 *
$p_1$	0.025*
$(1-p_1)$	0.975*
$p_2$	0.025*
$(1-p_2)$	0.975*
$\sigma_1$	0.50 *
$\gamma_2$	0.25 *
$P_s$	0.0001+
$(1-P_s)$	0.9999+
$\delta_q$	0.005*
$\delta_2$	0.03 *
$d$	0.016*
$q$	"0.25"
$(1-q)$	"0.75"
$\beta_e$	0.00001+

The data was obtained from different organizations such as the CDC (Center for Disease Control), American Lung Cancer Society, National Cancer Institute (NCI), and other non-profit and government agencies.

\* Estimated by the use of data, + Free Parameters = values assumed in order to try to make the model realistic, "-" Values that will be randomly changed to see the behavior of our model.



## 4.2 Numerical Solutions

We simulated our model in order to obtain the cases when  $R_0 < 1$  and  $R_0 > 1$ . Results show that when  $R_0 < 1$ , the population of susceptibles to the infection would eventually die out. This agrees with our analysis in section 3. The results of our simulations show that when  $R_0 > 1$ , an endemic stable state is established. The two simulations shown describe graphically what we have explained on the behavior of  $R_0$ .

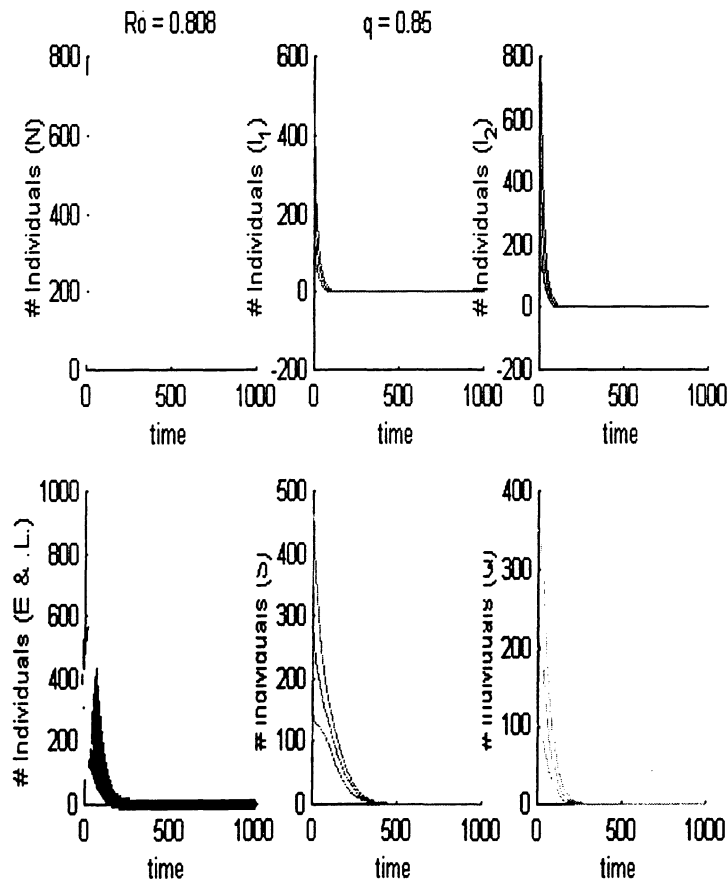


Figure 2: Graph shows the simulation for  $R_0 < 1$ . The smoke-free equilibrium is stable.

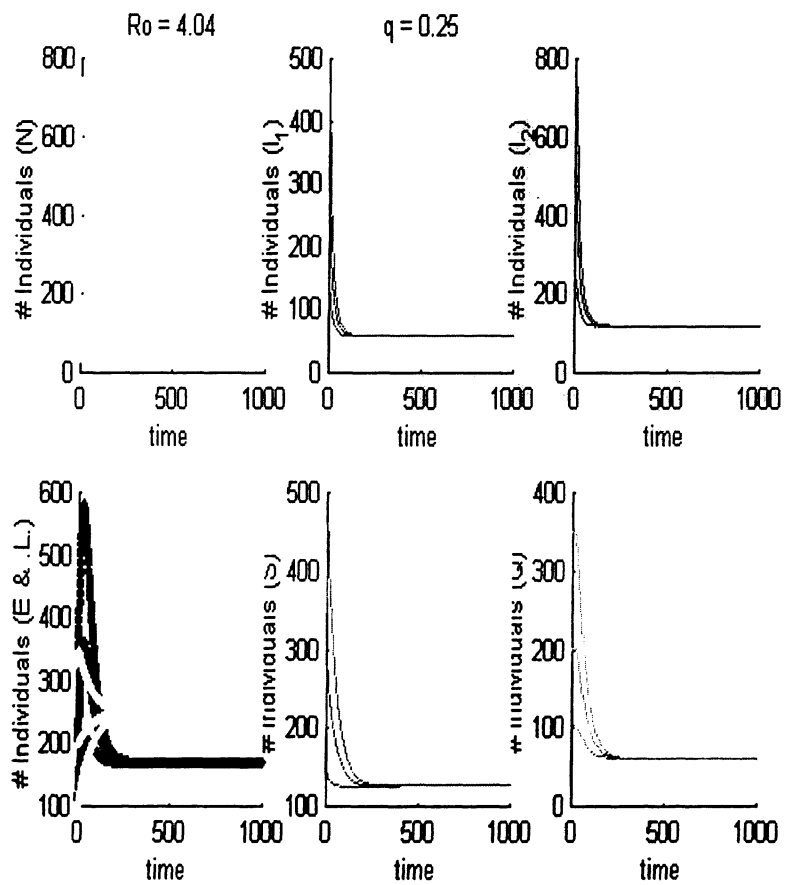


Figure 3: Graph shows the simulation for  $R_0 > 1$ . The endemic equilibrium is stable.

## 5 Results of our Deterministic Model

To obtain a realistic representation on the effects of smoking on a given population, we formulated a deterministic model. Simulations of our model were conducted using parameters and estimations from real life studies. We mainly varied two parameters to study the sensitivity of our results due to smoking within a population; these were  $q$ , the probability that incoming individuals would enter our educated class( $E$ ); and  $\beta$ , the transmission rate.

We produce several simulations showing the effects of smoking. Based on the analysis of  $R_0$  in the previous section and of our endemic-equilibrium point, we established that the epidemic of smoking will establish itself, provided that  $R_0 > 1$ . Otherwise, smoking will die out in our population. When  $R_0 > 1$ , we got  $q = 0.25$  and  $\beta = 2$ , which allowed us to obtain  $R_0 = 4.04$  (Figure 3). By looking at the seven classes, we can see that when  $R_0 > 1$ , then all seven classes will establish themselves. If we make  $q = 0.85$  and  $\beta = 2$ , then  $R_0 = 0.808$  (Figure 2). In this simulation, the population of smokers and ex-smokers eventually die out. The sum of the non-smoker class,  $N$ , and the educated class,  $E$ , reaches the approximate  $q$ -dependent equilibrium values. This simulation shows that for the given values, there will be no smoke-induced population.

When  $\sigma_1$  and  $\gamma_2$  consists of high values, that is when light and heavy smokers quit permanently and at a faster rate than likely smokers becoming smokers, then  $R_0$  should eventually be less than 1. If individuals from the light-smoking class quit at a higher rate than individuals that become heavy smokers, then we will be left with a smaller population of smokers in general. From the simulations, we see that if we make  $R_0 > 1$ , but close to 1, then the smokers ( $I_1$  and  $I_2$ ) will be a small population, but we will still have a very large portion of the total population in the temporary quitters ( $S$ ), meaning that individuals are still susceptible to smoking again. If we make  $\sigma_1$  high enough so that  $R_0 < 1$ , then the total population will be concentrated in the likely-smoker class ( $N$ ) and the educated population ( $E$ ).

Simulations where  $\gamma_2$  is varied can affect the values of  $R_0$  significantly, if we let  $\gamma_2 \rightarrow \infty$ , then the  $R_0$  equation could be less than 1. What this does to the equation is to eliminate the contribution of heavy smokers, but, since we still have the contribution of light-smokers, we can not necessarily say that the equation for  $R_0$  will be lower than 1 (Figure 6 and 7).

When we ran simulations varying  $\beta$ , we found out that if we made  $\beta$  high enough, then the smoking populations would establish themselves and the prevalence of smokers ( $\frac{I_1+I_2}{T}$ ) grows. Also, when  $\beta$  is a high value, the smokers will convert faster the likely smokers; then, our  $R_0$  and the risk of lung cancer increases (Figure 4). If we decrease  $\beta$  to a point where it is close enough to zero, then less individuals will start smoking due to peer pressure. Eventually,  $R_0$  could be less than 1 (Figure 5).

The parameters that we need to change in order to reduce the prevalence of smoking and lung cancer are  $q$ ,  $\beta$ ,  $\sigma_1$ , and  $\gamma_2$ . Ideally, one must concentrate on the most sensitive parameters, which are  $q$  and  $\beta$ . We say that because shown by the data of the parameters  $\sigma_1$ , and  $\gamma_2$ , it is much harder to affect individuals since there is a high percentage of smokers tht would like to quit, however, only 2.5 percent of those do it.

From our simulations, we observed that  $P_n$  (0.00001) does not have a big effect at the population level of lung cancer. For  $P_n$  to have a significant change in  $R_0$ , the value would have to change dramatically; however, the data indicates the opposite.

## 6 Conclusion

In our model the use of non-linear differential equation was crucial to study the dynamics of lung cancer at the population level caused by smoking and second-hand smoke. By building this population, we found an important aspect of mathematical biology,  $R_0$ , which controls the dynamics of our model.

On August 5, 2000, an article based on lung cancer was pulished in the *Ithaca Journal*, which came from a *British Journal of Medicine*. This article stated that if we decrease the education on non-smokers and concentrate on smokers, than the prevalence of a smoker developing lung cancer is low. However, using our model along with our simulations, we argue that when there is an increase of the number individuals that are educated, than their probability of becoming smokers decreases and eventually we will have a smoke-free population ( $R_0 < 1$ ). However, if  $R_0 > 1$ , then our population of light and heavy smokers will establish themselves. By changing  $\beta$ , we found that it had a significant effect on the number of individuals that were infected. But, the greatest difference ocurred where the value of  $q$  changed and when we educated a high number of individuals in our population.

In conclusion, the best way to lower the number of smokers and individuals who develop lung cancer is by increasing the number of individuals that are well-educated on the effect of smoking.

## 7 Future Work

Even though we considered the total population of smokers in our model, we can add to our conditions a number of variations. An age structure and ethnicity diversification can be added that will study and analyze the prevalence of lung cancer. This is due to the fact that smoking and its consequences are different if we take into account age, sex, and ethnicity. Also, studying a more realistic model that deals with the impact of smoking and the behavior it has on the prevalence of lung cancer. One example is studying certain brands of cigarettes. Also, we could build a model that would incorporate the recovery rates for lung cancer, meaning to add another class, a recovery class ( $R$ ), were the population of the lung cancer class ( $L$ ) could go. Looking into the development of lung cancers, we could take into consideration creating a model that looks at the effects of two types of lung cancers, since once an individual recovers from lung cancer the first time, Type 1, then they have a chance of getting a new type of lung cancer, Type 2. Finally, we could forward our research by looking at the effects of reducing the impact of peer pressure on likely new smokers, such as current smokers and the mass media.

## 8 Acknowledgement

This study was supported by the following institutions and grants: National Science Foundation (NSF Grant DMS-9977919); National Security Agency (NSA Grants MDA-904-00-1-0006); Presidential Faculty Fellowship Award (NSF Grant DEB) and the Presidential Mentoring Award (NSF Grant HRD) to Carlos Castillo-Chávez and the office of the provost of Cornell University; Intel Technology for Education 2000 Equipment Grant.

Thanks to our advisors Baojun Song and Carlos Castillo-Chávez, for your help and ideas. Also we want to thank Carlos Hernández, Steve Wirkus, Brisa Ney Sanchez and the MTBI program for allowing us to have the opportunity of doing this research. Finally thanks to MTBI students for your moral support.

## References

- [1] Brauer, Fred. *A Model for an SI Disease in Age-Structured Population*.
- [2] Brauer, Fred and Carlos Castillo-Chavez. *Methods and Principles in Mathematical Biology with Applications to Epidemiology Population Biology, and Resource Management*. (2000) pp. 271-294.
- [3] Cook, Allan R. *The New Cancer Sourcebook*. Omnigraphics (1996) pp. 6-10, 41-48, 251-282.
- [4] Gauderman, James W. and John L. Morrison. *Evidence for Age Genetic Relative Risks in Lung Cancer*. American Journal of Epidemiology. Volume 151, Number 1, January 1, 2000 pp.41-49.
- [5] Heesterbeek, Hans.  $R_0$ . pp.3-115.
- [6] United States Department of Health & Human Services. *SEER Cancer Statistics Review, 1973-1994*. Ed. Lynn A. Gloeckler Ries, M.S. pp 15-105.
- [7] Zhang, Tongxiao. *Role of Peer Pressure of Smoking: Smoking as an Epidemic*. Thesis (1999).
- [8] <http://www.aacr.org/5000/5001/5110g.html>. *Lung Cancer*. March 1998. pp. 1-5
- [9] <http://www.cancer.org/statistics/cff2000/tobacco.html>. *2000 Facts Figures: Tobacco Use* pp. 3-4.
- [10] [http://cancernet.nci.nih.gov/wyntk\\_pubs/lung.htm#5](http://cancernet.nci.nih.gov/wyntk_pubs/lung.htm#5) *Lung Cancer: Who's at Risk?* pp. 1-2.
- [11] <http://www.lungcheck.com/info/stats.htm>. *Lung Check-Statistical Facts*. pp. 1-3.
- [12] [http://rex.nci.nih.gov/NCI\\_Pub\\_Interface/raterisk/risks70.html](http://rex.nci.nih.gov/NCI_Pub_Interface/raterisk/risks70.html). *Risk Factors: Cigarette Smoking as a Cause of Cancer* pp.1-2.
- [13] [http://www.cdc.gov/tobacco/tab\\_3.htm](http://www.cdc.gov/tobacco/tab_3.htm). *CDC'S TIPS-Current Smoking Status Among Adults*. (1965-95) pp.1-3.
- [14] <http://www.cdc.gov/tobacco/mortali.htm>. *CDC'S TIPS - Cigarette Smoking-Related Mortality*.

- [15] <http://www.cdc.gov/tobacco/initfact.htm>. *CDC'S TIPS - Incidence of Initiation of Cigarette Smoking Among US Teens-Fact Sheet*. pp.1-2.
- [16] <http://www.cdc.gov/tobacco/adstat1.htm>. *CDC'S TIPS - Adult Prevalence Data*.
- [17] <http://www.cdc.gov/tobacco/adstat3.htm>. *CDC'S TIPS - Adult Prevalence Data*.
- [18] <http://www.cdc.gov/tobacco/adstat4.htm>. *CDC'S TIPS - Adult Former Smoker Prevalence Data*.

## 9 APPENDIX

In order to work on the simulations, we needed to create a program in MATLAB that was composed basically of the differential equations, the data we found, and the plotting of the graphs.

In MATLAB, we needed to build two programs in order to run the simulations.

Program 1

```
tspan=[0 1000];
x0=[500 200 200 200 200 200 200];
y0=[250 100 100 100 100 100 100];
z0=[750 350 350 350 350 350 350];
q = 0.25;
μ = 0.014;
β = 2;
δ1 = 0.015;
δ2 = 0.03;
δq = 0.01;
Λ = 14;
Ps = 0.0001;
Pn = 0.00001;
p1 = 0.025;
p2 = 0.025;
γ1 = 0.6;
γ2 = 0.25;
d = 0.016;
βe = .00001;

[t, x] = ode45('lung', tspan, x0, [], μ, β, δ1, δ2, Λ, Ps, Pn, p1, p2, γ1, σ1, γ2, δq, d, q, βe); [s, y] =
ode45('lung', tspan, y0, [], μ, β, δ1, δ2, Λ, Ps, Pn, p1, p2, γ1, σ1, γ2, δq, d, q, βe); [r, z] = ode45('lung', tspan, z0, [], μ, β, δ1, δ2, Λ, Ps, Pn, p1, p2, γ1, σ1, γ2, δq, d, q, βe); R0 = (1 - q) * (((1 - Pn) * β) / (γ1 + δ1 + σ1 + μ))
+ ((γ1 * (1 - Pn) * β) / ((γ1 + δ1 + σ1 + μ) * (γ2 + δ2 + μ)));
figure
subplot(231)
hold on
plot(t, x(:,1), 'c')
plot(s, y(:,1), 'b')
plot(r, z(:,1), 'm')
title(['R0 = ', num2str(R0)]);
```



```

xlabel('time')
ylabel('# Individuals ( $N$ )')
hold off
subplot(232)
hold on
plot(t,x(:,2),'r')
plot(r,z(:,2),'g')
plot(s,y(:,2),'b')
title(['q = ',num2str(q)]);
xlabel('time')
ylabel('# Individuals ( $I_1$ )')
hold off
subplot(233)
hold on
plot(t,x(:,3),'r')
plot(s,y(:,3),'b')
plot(r,z(:,3),'g')
xlabel('time')
ylabel('# Individuals ( $I_2$ )')
title([' $\beta$  = ',num2str( $\beta$ )]);
hold off
subplot(236)
hold on
plot(t,x(:,4),'g')
plot(s,y(:,4),'m')
plot(r,z(:,4),'y')
xlabel('time')
ylabel('# Individuals ( $Q$ )')
hold off
subplot(235)
hold on
plot(t,x(:,5),'g')
plot(s,y(:,5),'m')
plot(r,z(:,5),'b')
xlabel('time')
ylabel('# Individuals ( $S$ )')
hold off
subplot(234)

```

```

hold on
plot(t,x(:,6),'g.')
plot(s,y(:,6),'y.')
plot(r,z(:,6),'k.')
plot(t,x(:,7),'r')
plot(s,y(:,7),'m')
plot(r,z(:,7),'b')
xlabel('time')
ylabel('# Individuals (E & .L.)')
hold off

```

## Program 2

```

function
dx=lung(t,x,flag,mu,beta,delta1,delta2,Lambda,P_s,P_n,p1,p2,gamma1,sigma1,gamma2,delta_q,d,q,beta_e)

N = x(1); I1 = x(2); I2 = x(3); Q = x(4); S = x(5); L = x(6); E = x(7);
T = N + I1 + I2 + Q + S + L + E;
eq1 = (1 - q) * Lambda - beta * N * (I1 + I2) / T - mu * N;
eq2 = (1 - P_n) * beta * N * (I1 + I2) / T + (1 - P_s) * beta * S * (I1 + I2) / T - (sigma1 + gamma1 + delta1 + mu) * I1;
eq3 = gamma1 * I1 - (gamma2 + delta2 + mu) * I2;
eq4 = p2 * gamma2 * I2 + p1 * sigma1 * I1 - (delta_q + mu) * Q;
eq5 = (1 - p1) * sigma1 * I1 + (1 - p2) * gamma2 * I2 - beta * S * (I1 + I2) / T - mu * S;
eq6 = P_n * beta * N * (I1 + I2) / T + P_s * beta * S * (I1 + I2) / T + beta_e * E * (I1 + I2) / T + delta1 * I1 + delta2 * I2;
eq7 = q * Lambda - beta_e * E * (I1 + I2) / T - mu * E;

dx = [eq1; eq2; eq3; eq4; eq5; eq6; eq7];

```

In this section, we will show some other simulations that will explain the behavior of our model if we vary some other parameters that were supposed to changed the value of  $R_0$  significantly.

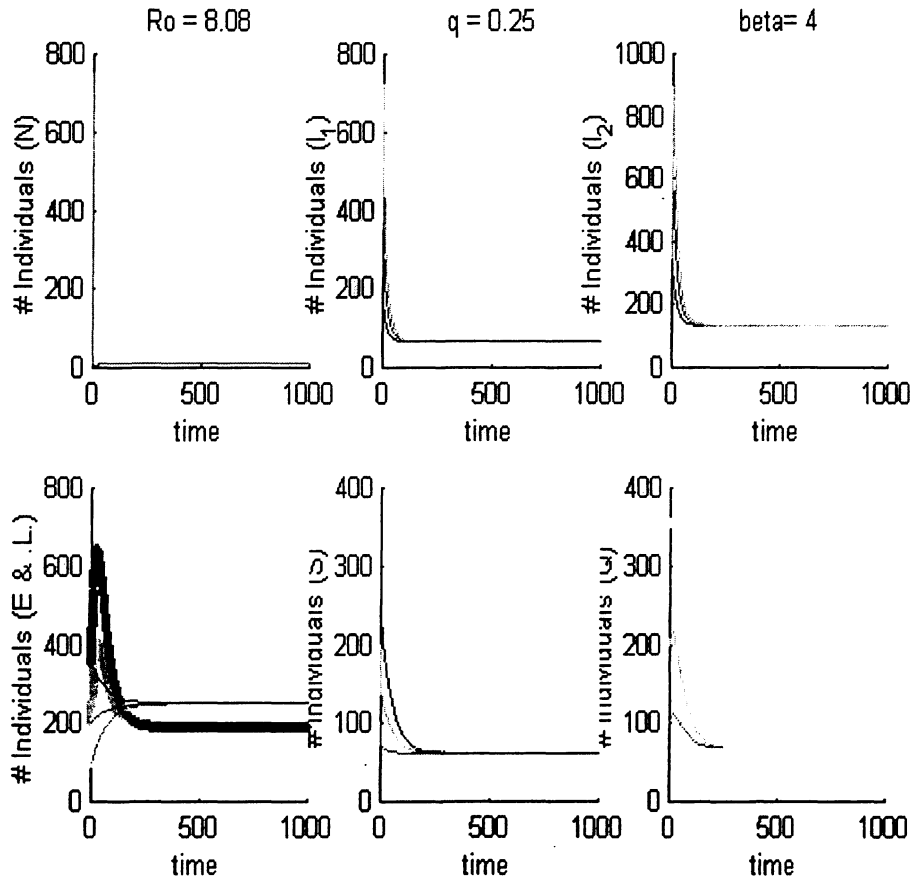


Figure 4: This graph is with  $\beta = 4$ .

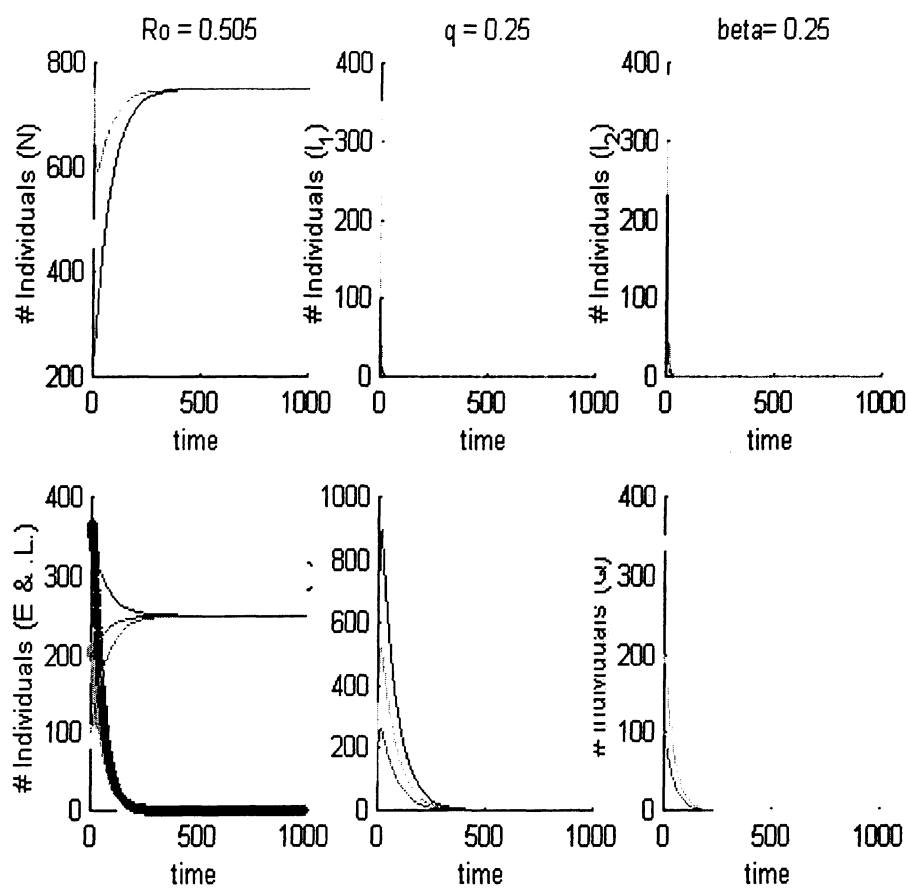


Figure 5: This graph is with  $\beta = .25$

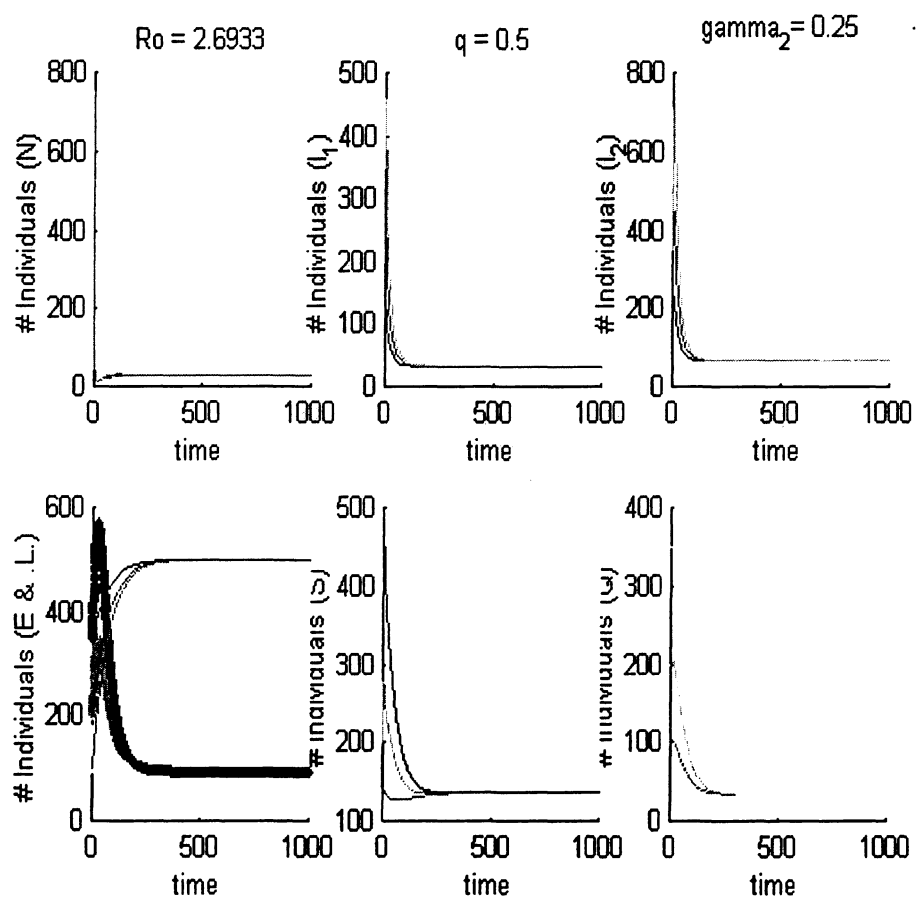


Figure 6: This graph is with  $\gamma_2 = 0.25$ .

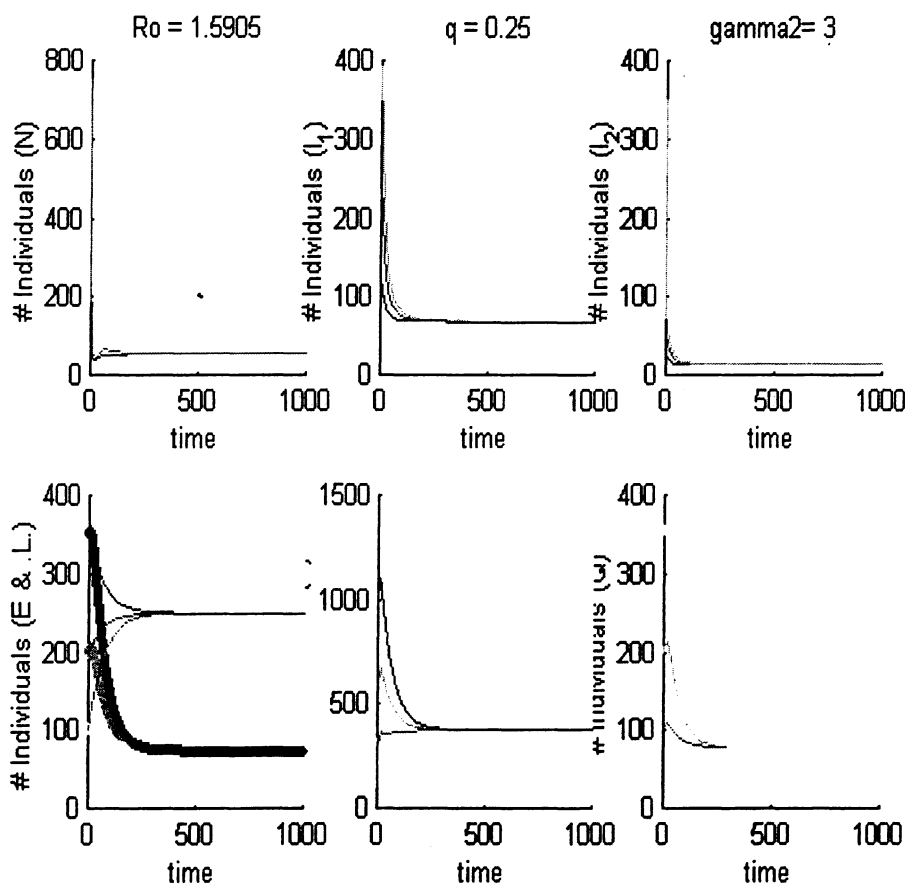


Figure 7: This graph is with  $\gamma_2 = 3$ .

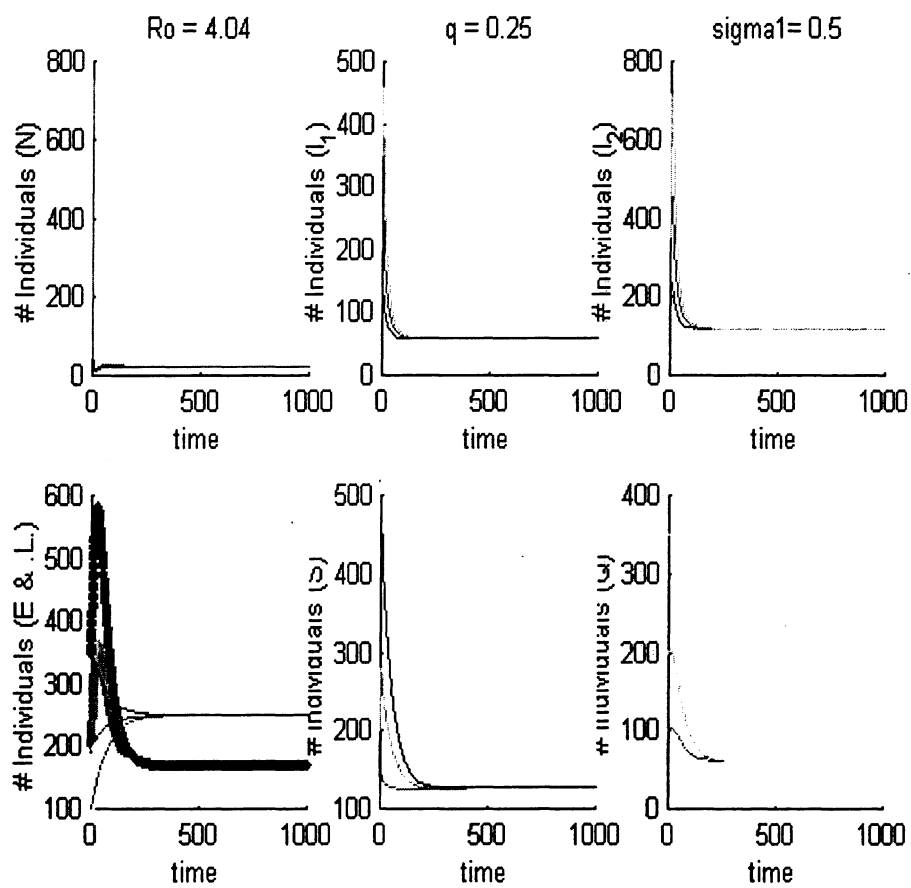


Figure 8: This graph is with  $\sigma_1 = 0.5$ .

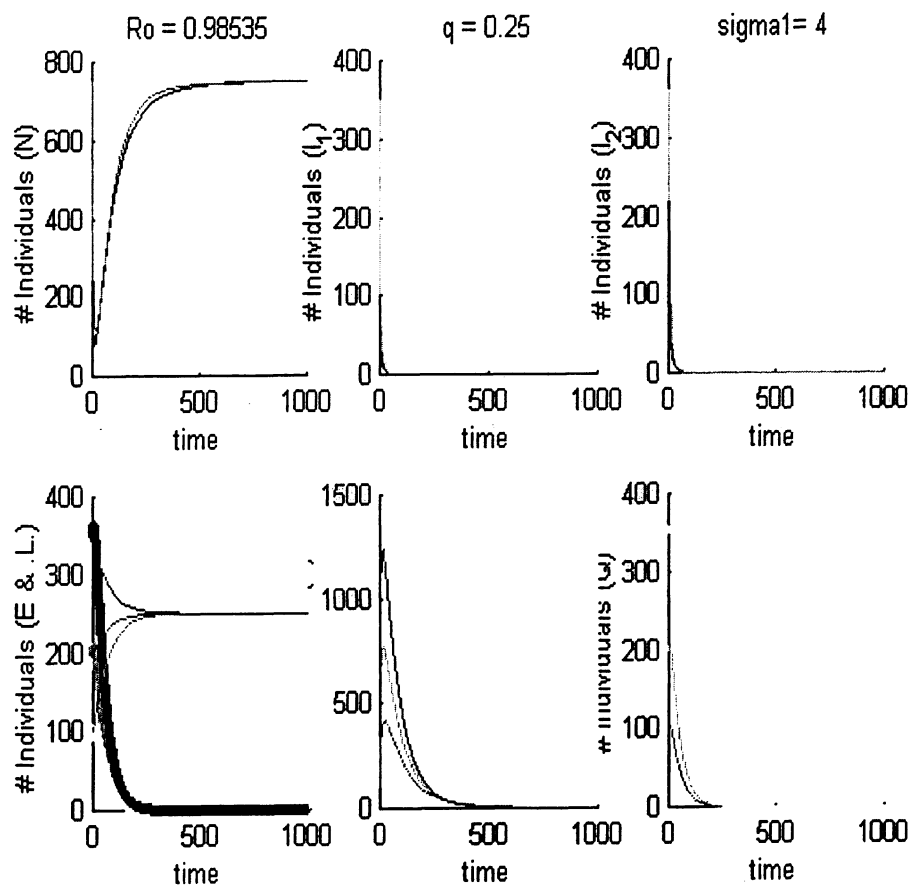


Figure 9: This graph is with  $\sigma_1 = 4$ .